Combining Newton interpolation and deep learning for image classification

Zhang, Yongfeng; Shang, Changjing

Published in:
Electronics Letters

DOI:
10.1049/el.2014.3223

Publication date:
2015

Citation for published version (APA):
Zhang, Y., & Shang, C. (2015). Combining Newton interpolation and deep learning for image classification. *Electronics Letters*, *51*(1), 40-42. https://doi.org/10.1049/el.2014.3223

Automatic Segmentation of Martian Rock Images

C. Gui and C. Shang

Accurate segmentation of Mars rock images forms an essential step towards automatic detection of rocks on Martian surface, which is of great interest to geologists and planetary scientists. This letter introduces a new approach to segmenting Mars images, supported by improved Otsu's thresholding and Canny edge detection that are utilised to help identify potential target regions and their space relations, respectively. Closed contours of rocks are determined through template dilatation edge linking. Experimental results with real images captured by the NASA Mars Exploration Rover, which are of different illuminations, spectral bands and scenes are shown, demonstrating that the proposed approach is consistent with human perception and outperforms typical existing methods over a range of performance criteria.

Rocks are one of the observed main features exposed on Martian surface. They provide rich information for the study of planetary geology, and offer good tie points for vision-based rover localisation and navigation. Therefore, automatic rock detection is invaluable for Mars exploration. A key step to ensure such detection is accurate segmentation of the Martian rock images.

There have been a number of techniques developed for segmenting Mars images. Shang and Barnes [1] constructed Mars image classifiers using fuzzy-rough feature selection combined with support vector machines. However, as the underlying colour images are currently manually produced by scientists on Earth, their method is impractical for use on-board. Gor et al. [2] proposed a rock detection approach where large rocks are identified using range data through ground plane fitting and height information, whilst edge-flow segmentation and image intensity are exploited for small rocks. This method requires the introduction of image scale as a control parameter and the range data is assumed from stereo imagery. Castano [3] reported a mechanism for determining the closed contours of rocks by combining an edge-based rock detection algorithm and multi-scale image pyramids. This method may not be sufficiently efficient in dealing with images where intensity differences exist between rocks and their surrounding soil. In addition, Fink adopted the conventional k-means clustering method to build clusters of rocks [4], and Pugh and Barnes [5] put forward a similar region-growing technique for rock segmentation.

This letter proposes an unsupervised approach to tackle the problem of segmentation rock regions from Martian images. A tri-level thresholding algorithm that extends the original Otsu's method [6] is employed to address the issue of background segmentation. Edges determined by the popular Canny method are utilised to merge sub-regions returned by Otsu thresholding, thereby identifying the rock boundaries as much as possible. Template dilatation edge linking [7] is subsequently employed to eventually detect the closed contour of each rock in a given image.

Tri-level Thresholding: For the present work, it is not just to tell rocks as foreground regions from the background, but it is also important to detect rock shadows. Therefore, the conventional Otsu's binarisation [6] cannot applied directly as it partitions an image into two classes only. To meet the current need, the bi-level thresholding method is herein extended to map image pixels onto three classes, with two thresholds. The improved algorithm is outlined below.

In traditional bi-level thresholding, an image is interpreted as a 2D greyscale intensity function, involving *N* pixel values with grey levels ranging from *0* to *G*, where *G* is typically 255 if 8-bit quantisaition is assumed. Let the number of pixels with grey level *i* be f_i, and the probability of the occurrence of grey level *i* in an image be approximated by the frequency:

$$p_i = f_i/N \tag{1}$$

Otsu's binarisation mechanism works by applying an exhaustive search that minimises intra-class variance to derive an optimal grey level threshold t, in order to split a given image into two groups of grey levels, say, C_1 and C_2. Here, without losing generality, C_1 represents the background and C_2 represents the foreground. The intra-class variance is defined as a weighted sum of variances of the two classes:

$$\sigma^2_{intra-class} = \omega_1(t)\sigma^2_1(t) + \omega_2(t)\sigma^2_2(t) \tag{2}$$

where weights $\omega_i(t), i = 1, 2$ are the probabilities of the two classes C_1 and C_2, separated by the (to-be-sought) threshold t and $\sigma_i, i = 1, 2$ are the variances of these classes, and the class probabilities are computed by

$$\omega_1(t) = \sum_{i=0}^{t} p_i, \quad \omega_2(t) = \sum_{i=t+1}^{G} p_i \tag{3}$$

It has proven that minimising the above intra-class variance is equivalent to maximising the inter-class variance:

$$\sigma^2_{inter-class} = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \tag{4}$$

where the class means are defined by

$$\mu_1(t) = \sum_{i=0}^{t} x_i p_i/\omega_1(t), \quad \mu_2(t) = \sum_{i=t+1}^{G} x_i p_i/\omega_2(t) \tag{5}$$

where x_i is the value at the centre of the ith histogram bin. Given this, the desired binarisation threshold is the t that maximises:

$$\sigma^2_{inter-class}(t), \quad 0 \le t \le G \tag{6}$$

When the foreground C_2 is required to be further split into another two classes of shadows and rocks, the above process is repeated within regions that collectively represent C_2 in the same manner.

Regions Merging: Through the use of tri-level thresholding, a number of areas within an image may be regarded as parts of rocks. Empirically, it is known that certain areas belong to one single rockInterpreting such areas as individual rocks may obviously result in serious over-segmentation. To address this problem, spatial details are generally taken into account in order to establish relationships between adjacent areas.

Regions merging is proposed here to combine areas which are potential portions of the same rock by consolidating their individual edges into an identical common edge. Canny's edge detector is adopted here to help accomplish this task. Post-processing of the detected edges is required, however, with the following edge cleaning heuristics implemented in the present work:

- Connecting endpoints. Because a vast majority of detected edges with the Canny's methods are typically unable to generate a closed contour (relating to a potential target of interest), those endpoints off one pixel are connected with a straight line.
- Breaking edges. Because a solid contour of a target of potential interest is assumed to be of no branched edges, those edges with points that form a Y-junction are forced to be broken.

Let $R_i, i = 1, 2, ..., n$ denote the areas returned by tri-level thresholding, $E_j, j = 1, 2, ..., m$ represent the edges returned by the heuristic post-precessing, and L_i express the collection of those edges E_k, each of which forms part of the contour of the area R_i:

$$L_i = \{E_u | R_i \cap E_u \neq \Phi, u \in \{1, 2, ..., m\}\} \tag{7}$$

The core of the proposed method for regions merging is a recursion over $s, 1 \le s \le n$, leading to revised areas and their associated edge sets $\{E_t | O_s \cap E_t \neq \Phi\}, 1 \le t \le m$:

$$O_s = \begin{cases} R_i \cup R_j, & if \quad L_i \cap L_j \neq \Phi, \quad i \neq j \\ R_i, & otherwise \end{cases} \tag{8}$$

$$E_t = \begin{cases} E_u \cap E_v, & if \quad L_i \cap L_j \neq \Phi, \quad i \neq j \\ E_u, & otherwise \end{cases} \tag{9}$$

where $L_j = \{E_v | R_j \cap E_v \neq \Phi, v \in \{1, 2, ..., m\}\}$.

In summary, each recursion of the merging process results in a revised $L_s = \{E_t | O_s \cap E_t \neq \Phi, t \in \{1, 2, ..., m\}\}$. This method is efficient as it involves set operations only. The final output of the merging process is a set of merged areas $\{O_k | k \in \{1, 2, ..., v\}, 1 \le v \le n\}$. Note that in the unlikely event where $v = n$ the result implies that there is no merge taking place, meaning that all areas originally returned by tri-level thresholding are deemed to be individual rocks.

Template Dilatation Edge Linking: The merging of the areas and edges is to ensure individual regions are bounded as much as possible. Unfortunately, it is not guaranteed that all target areas are allocated with closed contours (along the border of each rock), many rocks may still include blurred edges. Thus, the template dilatation edge linking method as proposed in [7] is applied to close contours of all detected rocks in a given image.

Briefly, this algorithm works as follows: 1) Identify all those areas which have an unclosed boundary. 2) For each unclosed area, choose a random pixel as the start point on the boundary. 3) Construct an $m \times m$ template centred on the start pixel (empirically, $m = 3$). 4) Check if there are any pixels belonging to other classes in the template; if so, (a) If these pixels are in the same class, then sequentially connect them up using a straight line, (b) If they belong to different classes, then connect the nearest two pixels of the different classes using a straight line; then go to step 6; else, repeat until all pixels have been visited. 5) Dilate the current template, by setting $m = m + 2$ and go to step 4. 6) Choose the next neighbouring pixel point relative to the current point on the boundary. 7) If the pixel point is the end of all sequential traversal pixels on the boundary, then go to step 2; if no unclosed area remains, then the algorithm ends; or else, go to step 3.

Experimental Results: The proposed approach has been implemented and run on a set of six (large scale) representative real Martian images, taken from two datasets (Navcam imagery and Pancam imagery respectively) acquired by the NASA Mars Exploration Rover (MER), as listed in Table 1. Importantly, these images reflect distinct illuminations, spectral bands and scenes. Both qualitative and quantitative comparisons with three other state-of-the-art methods are carried out.

Table 1: Selected images for experimentation (Spirit Rover)

Image Name	Image NO.	Local Solar Sol/Time	Spectral Band	Instrument
2n136853953ilf4000p1977l0m1	1.png	118/16:04:42	-	NAVCAM
1p163700052edn5000p2384l2m1	2.png	400/14:43:25	753NM	PAMCAM
2p130811027edn1000p2421l6m1	3.png	50/14:00:54	483NM	PAMCAM
2p130974937eff1100p2568l7m1	4.png	52/10:20:08	440NM	PAMCAM
2p162142844esfa600p2558l5m1	5.png	403/13:42:44	535NM	PAMCAM
2p179724220esfaeghp2570l2c1	6.png	601/15:40:32	753NM	PAMCAM

For qualitative comparison, the segmentation results of the present work and those obtained by running the Otsu's, Pugh's [5] and Fink's [4] methods are shown in Fig. 1. These results show that the segmentation effects of the proposed approach are more consistent with the human visual perception than the others.

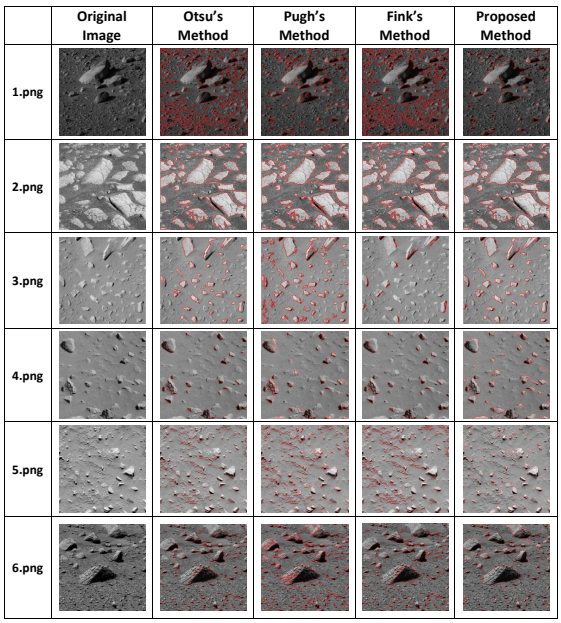

Fig. 1. Qualitative comparison

Quantitatively, the proposed and other approaches are assessed and compared with respect to the following three criteria: precision, recall and misclassification rate. The performance is evaluated between the grey-level images, of Fig. 1, and the corresponding ground-truth images which include a total of 128 rocks that have been identified manually. In particular, precision denotes the ratio of the number of pixels within those detected targets which are indeed rocks over that of those which are detected as rocks in an image; recall represents the fraction of the number of pixels within those detected targets which are indeed rocks over that of those which are true rocks; and misclassification rate shows the proportion of the total of foreground pixels being mistakenly allocated to the background and the background pixels being falsely attributed to the foreground across the whole image.

Tables 2-4 summarise the segmentation performances of the four approaches under comparison. In terms of precision, for three images (1.png, 5.png and 6.png) the best accuracy is achieved by the proposed work, and the proposed method significantly outperforms the others in these cases. Although the best for the other three is obtained interestingly, by one of the other three method, each for one image, the results of using the proposed are not far aware from such best ones. For recall, the work present here beat the others in five out of the six cases, whilst for the other image (3.png) its performance is the second best, just after Pugh's method. Importantly, with regard to the criterion of misclassification rate, the proposed method completely outperforms the other approaches. Note that even under the criteria of precision and recall where the proposed method may not always beat the other techniques tested for certain images, the average accuracy as reflected in the right-most column of each of these result tables, the present work attains the best performance for all images.

Table 2: Comparison: Precision

Method	1.png	2.png	3.png	4.png	5.png	6.png	AVG
Otsu's	0.261	0.955	0.712	0.942	0.314	0.679	0.644
Pugh's	0.739	0.95	0.593	0.975	0.446	0.703	0.734
Fink's	0.246	0.948	0.846	0.967	0.247	0.663	0.653
Proposed	0.824	0.936	0.78	0.881	0.589	0.81	0.803

Table 3: Comparison: Recall

Method	1.png	2.png	3.png	4.png	5.png	6.png	AVG
Otsu's	0.533	0.808	0.532	0.298	0.297	0.433	0.484
Pugh's	0.785	0.826	0.791	0.387	0.722	0.708	0.703
Fink's	0.534	0.82	0.133	0.289	0.312	0.434	0.42
Proposed	0.871	0.868	0.703	0.784	0.785	0.857	0.811

Table 4: Comparison: Misclassification rate

Method	1.png	2.png	3.png	4.png	5.png	6.png	AVG
Otsu's	0.332	0.091	0.09	0.109	0.108	0.155	0.148
Pugh's	0.083	0.089	0.099	0.094	0.094	0.119	0.096
Fink's	0.353	0.089	0.117	0.109	0.131	0.158	0.16
Proposed	0.053	0.076	0.065	0.049	0.061	0.069	0.062

Conclusion: This letter has presented an unsupervised method for image segmentation for complicated Mars terrain images. Both qualitative and quantitative comparisons have been carried out with a number of existing approaches, using large scale real Mars images, against multiple criteria, including precision, recall and misclassification rate. The results demonstrate the general superior performance

C. Gui and C. Shang (*Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK*)

E-mail: cns@aber.ac.uk

References

1 Shang, C. and Barnes, D.: 'Fuzzy-rough feature selection aided support vector machines for Mars image classification', *Comput. Vis. Image Und.*, 2013, **117**, pp. 202-213
2 Gor, V. et al.: 'Autonomous rock detection for Mars terrain', *Proc. AIAA*, 2001, pp. 1-14
3 Castano, R.: 'OASIS: Onboard autonomous science investigation system for opportunistic rover science', *J. Field Robot.*, 2007, **24**, 379-397
4 Fink, W. and Datta, A.: 'Automated global feature analyzer: A driver for tier-scalable reconnaissance', *Proc. IEEE Aerospace*, 2008, pp. 1-12
5 Pugh, S. and Barnes, D.: 'Autonomous sample selection and acquisition for planetary exploration', *Proc. European Planetary Science*, 2007
6 Otsu, N.: 'A threshold selection method from gray level histogram', *IEEE T. Syst. Man Cy.*, 1979, **9**, 62-66
7 Gui, C., Barnes, D. and Pan, L.: 'A SIFT-based method for matching desired keypoints on Mars rock targets', *Proc. AIRA Space*, 2012